(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 December 2000 (21.12.2000)

PCT

(10) International Publication Number
WO 00/77031 A2

(51) International Patent Classification[7]: C07K 7/00

(21) International Application Number: PCT/CA00/00711

(22) International Filing Date: 15 June 2000 (15.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/139,263 15 June 1999 (15.06.1999) US

(71) Applicant *(for all designated States except US)*: **UNIVERSITY HEALTH NETWORK** [CA/CA]; Room 7-504, 610 University Avenue, Toronto, Ontario M5G 2M9 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **GARIEPY, Jean** [CA/CA]; 29 Elmsthorpe Avenue, Toronto, Ontario M5P 2L5 (CA). **YANG, Shaoxian** [CA/CA]; 309-36 Thorncliffe Park Drive, Toronto, Ontario M4H 1J8 (CA).

(74) Agent: **BERESKIN & PARR**; 40th floor, 40 King Street West, Toronto, Ontario M5H 3Y2 (CA).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/77031 A2

(54) Title: MUC1 LIGANDS

(57) Abstract: Novel ligands that bind to MUC1 are disclosed. The ligands were isolated using an improved phage display technique using MUC1 tandem repeat as a target. Uses of the ligand to detect, monitor or treat cancer as well as to prepare antibodies is also described.

Title: MUC1 Ligands

FIELD OF THE INVENTION

The present invention relates to ligands that bind to MUC1 and uses of the ligands to diagnose, monitor and treat cancer.

BACKGROUND OF THE INVENTION

Mucins are high molecular weight glycoproteins which are expressed on the surface of normal and cancer tissues. The epithelial cell mucin encoded by the MUC1 gene, is commonly expressed by adenocarcinomas of the pancreas, breast, ovary as well as by several other tumors of ductal epithelial cell origin. The protein core of mucins contains a variable number of tandem repeats (VNTRs) rich in O-glycosylation sites. The aberrant glycosylation of mucin tandem repeats by malignant cells results in the display of unique tumor-associated structures on their surface (1-4). In particular, mammary epithelial cells express mucins harbouring 40 to 80 copies of the MUC1 tandem repeat, a 20-amino acid long domain with the sequence Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala (5-8). All threonines and serines present in the MUC1 tandem repeat are normally O-glycosylated in the context of milk-derived mucin (9), a process initiated by at least four distinct human UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases (10,11). A comprehensive analysis of 56 monoclonal antibodies that bind to the tumor-associated MUC1 mucin has revealed that a majority of them recognize peptide epitopes within the sequence Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro (ref.12 and references thereafter) suggesting that the threonine residue within this sequence is underglycosylated in malignant cells exposing this segment of the tandem repeat. In addition, mucin-specific IgMs have been detected in sera from patients with breast and ovarian cancers and have been shown to recognize this epitope (13-15).

MUC1 has clinical relevance as it is found on the surface of many tumors and is a useful diagnostic and therapeutic tool. In addition,

MUC1 is shed from the surface of many cancers and serum concentrations may be useful for detecting certain cancers and monitoring the response to cancer therapy. Expression of MUC1 is also used for micrometastatic tumor cell detection in patients with solid tumors such as breast cancer. MUC1 is also a potential target for tumor immunotherapy.

To date, clinical strategies aimed at targeting MUC1-expressing tumor cells have been limited to the use of monoclonal antibodies and related fragments in guided immunotherapies (16-19) and at developing vaccination approaches based on the MUC1 tandem repeat (20-22). Small MUC1-binding peptides would offer clear advantages over antibodies (12) and their single chain variants (23), in terms of structural simplicity, penetration into solid tumors (24-26) and their lack of immunogenicity.

In view of the foregoing, there is a need in the art to develop small ligands that bind MUC1 which would be useful diagnostic and therapeutic tools.

SUMMARY OF THE INVENTION

The present inventors have isolated several novel peptide ligands that bind to MUC1. The ligands were isolated using an improved phage display technique using MUC1 tandem repeats as a target.

Accordingly, the present invention provides isolated and purified ligands that bind to MUC1.

In one embodiment, the ligand comprises the sequence X_1-X_2-X_3 wherein:

X_1 is Pro, Thr or Ser;

X_2 is any amino acid; and

X_3 is Pro, Ala, Lys, His, Thr, Asn, Tyr or Phe.

The present invention further provides a method of isolating a ligand that binds to MUC1 comprising:

(a) providing a MUC1 target molecule wherein the MUC1 target molecule comprises at least 3 tandem repeats of the sequence PDTRPAPGSTAPPAHGVTSA;

(b) contacting the MUC1 target molecule with a peptide or small molecule library, under conditions to allow the MUC1 target molecule to bind to any MUC1 ligands in the library, to prepare a test sample;

(c) panning the test sample under high stringency conditions; and

(d) isolating the MUC1 ligands bound to the MUC1 in the test sample.

The present invention also provides the use of the ligands that bind MUC1 to detect, monitor or treat cancer as well as to prepare antibodies.

Other features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples while indicating preferred embodiments of the invention are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention will now be described in relation to the drawings in which:

Figure 1A-C are graphs showing the binding of biotinylated MUC1-Lp1, MUC1-Lp2 and MUC1-Lp3 to MUC1 octapeptides synthesized on plastic pins was measured at 450 nm using a streptavidin-HRP conjugate. Each bar represents the average absorbance readings calculated for assays performed on six copies of each MUC1 octapeptide.

Figure 2A-C are graphs showing the inhibition by MUC1 ligand peptides of anti-MUC1 mAb BC2 binding to the MUC1 tandem repeat. The mAb BC2 was dispensed into MUC1/3TR-coated wells in the presence or absence of MUC1 peptide ligands. The term *mAb BC2 % relative binding* represents the ratio, expressed as a percentage, of absorbance values between wells treated or not with increasing concentrations of a MUC1

ligand. Figure 2A, TMGFTAPRFPHY (MUC1-Lp1), ◆; TMGFTA, ❑; FTAPRF, Δ; PRFPHY, ◊; Figure 2B, SWWPFPPQPDPA (MUC1-Lp2), ◆; SWWPFP, ❑; PFPPQP, Δ; PQPDPA, ◊; Figure 2C, YTKSDPLKLLES (MUC1-Lp3), ◆; YTKSDP, ❑; SDPLKL, Δ; LKLLES, ◊. The peptide YTVTNSWTWWSPLQQA (■) does not bind to the MUC1 repeat and was used as a negative control.

Figure 3A-C are graphs showing the binding of radiolabelled MUC1 ligand peptides to cells. Curves depict the total (■, ❑) and non-specific (▲, Δ) binding of ^{125}I-labelled MUC1 ligands to either MUC1$^+$ T47D (■,▲) and MUC1$^-$ Daudi cells (❒,Δ). Figure 3A, binding curves for ^{125}I-labelled MUC1-Lp1; Figure 3B, binding curves for ^{125}I-labelled MUC1-Lp2; Figure 3C, binding curves for ^{125}I-labelled MUC1-Lp3.

Figure 4A and B are graphs showing competitive displacement of radiolabelled peptide ligands from T47D cells with unlabelled MUC1 ligands. Figure 4A, Displacement curves for ^{125}I-labelled MUC1-Lp1 (❑), MUC1-Lp2 (Δ) and MUC1-Lp3 (▲) in the presence of increasing concentrations of their respective unlabelled ligand. Figure 4B, Displacement curves for ^{125}I-labelled MUC1-Lp1 with either unlabelled MUC1-Lp1 (❑), MUC1-Lp2 (Δ) or MUC1-Lp3 (▲).

DETAILED DESCRIPTION OF THE INVENTION

The following standard one letter and three letter abbreviations for the amino acid residues may be used throughout the specification: A, Ala - alanine; R, Arg - Arginine; N, Asn - Asparagine; D, Asp - Aspartic acid; C, Cys - Cysteine; Q, Gln - Glutamine; E, Glu - Glutamic acid; G, Gly - Glycine; H, His - Histidine; I, Ile - Isoleucine; L, Leu - Leucine; K, Lys - Lysine; M, Met - Methionine; F, Phe - Phenyalanine; P, Pro - Proline; S, Ser - Serine; T, Thr - Threonine; W, Trp - Tryptophan; Y, Tyr - Tyrosine; and V, Val - Valine;

1. Ligands of the Invention

As hereinbefore mentioned, the present inventors have isolated several novel peptide ligands that bind to MUC1. Accordingly, the

present invention provides isolated and purified ligands that bind to MUC1.

In order to isolate MUC1 ligands with high specificity for MUC1, an improved screening assay was developed by the inventors which is discussed in greater detail below. In summary, a 100-amino acid long peptide sequence corresponding to 5 consecutive 20 amino acid long tandem repeats of the MUC1 mucin was chemically synthesized and used as a target in a solid-phase screening assay to probe a phage display library. The 20 amino acid long tandem repeat had the following sequence PDTRPAPGSTAPPAHGVTSA (SEQ.ID.NO.:1). The MUC1 target was directly applied to plates and contacted with a phage display library expressing random 12 amino acid sequences. The phage that bound to the MUC1 using high stringency conditions (i.e. high salt) were selected. Fifteen distinct, 12-residue long peptide sequences were identified after several rounds of phage planning which are shown in Table 1.

Using the improved screening assay, the following ligands were isolated: TMGFTAPRFPHY (SEQ.ID.NO.:2), SWWPFPPQPDPA (SEQ.ID.NO.:3), YTKSDPLKLLES (SEQ.ID.NO.:4), VVPVHWSRGVVL (SEQ.ID.NO.:5), HIPVAALAPRMT (SEQ.ID.NO.:6), LGLQPPTSALDP (SEQ.ID.NO.:7), TPAFSPLPTDLL (SEQ.ID.NO.:8), ELNTHLATNVFT (SEQ.ID.NO.:9), IDVHSINFLATL (SEQ.ID.NO.:10), THPWSLKSTSFF (SEQ.ID.NO.:11), YITPYAHLAGGN (SEQ.ID.NO.:12), SLPIPSHARLQN (SEQ.ID.NO.:13), YLPYATLSQNSH (SEQ.ID.NO.:14), WHIPPNIGRTFS (SEQ.ID.NO.:15) and TSNPHTRHYYPI (SEQ.ID.NO.:16).

The above MUC1 ligands isolated by the inventors' improved method share a common motif, X_1-X_2-X_3 wherein:

X_1 is Pro, Thr or Ser;

X_2 is any amino acid; and

X_3 is Pro, Ala, Lys, His, Thr, Asn, Tyr or Phe.

Preferably, X_1 is Pro or Ser; X_2 is Leu, Ile, Tyr, Ala, Val, Phe, Pro, Met, His, Arg, Gln, Asp or Glu and X_3 is Pro or Phe.

In one embodiment, the MUC1 ligand has the sequence TMGFTAPRFPHY (SEQ.ID.NO.:2) which is also referred to as MUC1-Lp1 herein. In order to determine the MUC1 binding site of MUC1-Lp1, the inventors synthesized hexapeptides covering the MUC1-Lp1 sequence and found that the sequence FTAPRF (SEQ.ID.NO.:17) and PRFPHY (SEQ.ID.NO.:18) were also able to bind MUC1. Both of these sequences share the common tripeptide PRF (SEQ.ID.NO.:19) consistent with the common motif X_1-X_2-X_3 described above. Accordingly, the present invention includes a MUC1 ligand comprising the sequence PRF, which includes the sequences FTAPRF and PRFPHY.

In another embodiment, the MUC1 ligand has the sequence SWWPFPPQPDPA (SEQ.ID.NO.:3) which is also referred to as MUC1-Lp2 herein. By synthesizing peptides spanning MUC1-Lp2, the inventors determined that the peptide SWWPFP (SEQ.ID.NO.:20) was a useful MUC1 ligand. The inventors further determined that peptides containing the motif P-X-P (SEQ.ID.NO.:21) such as Pro-Gln-Pro (SEQ.ID.NO.:22); Pro-Asp-Pro (SEQ.ID.NO.:23); and Pro-Phe-Pro (SEQ.ID.NO.:24) could also bind MUC1 which is consistent with the formula X_1-X_2-X_3 above. Accordingly, the present invention includes MUC1 ligand comprising the sequence P-X-P, wherein X can be any amino acid.

In a further embodiment, the MUC1 ligand has the sequence YTKSDPLKLLES (SEQ.ID.NO.:4) which is referred to as MUC1-Lp3 herein. By synthesizing peptides spanning MUC1-Lp3, the inventors determined that the sequences SDPLKL (SEQ.ID.NO.:25) and YTKSDP (SEQ.ID.NO.:26) could also bind MUC1. These sequences share the motif Ser-Asp-Pro (SEQ.ID.NO.:27) which is consistent with the formula X_1-X_2-X_3 above. Accordingly, the present invention includes a MUC1 ligand comprising the sequence SDP, which includes the sequences SDPLKL and YTKSDP.

The term "MUC1 ligand" as used herein means a peptide that can bind MUC1. The term includes peptides that are of a formula or sequence as defined above and fragments, analogs and derivatives of the

peptides which maintain the ability to bind MUC1. The term also includes peptides isolated according to the method of the invention.

The term "analog" includes any peptide having an amino acid residue sequence substantially identical to the sequence of the MUC1 ligands described herein in which one or more residues have been conservatively substituted with a functionally similar residue and which displays the ability to mimic a MUC1 ligand. Examples of conservative substitutions include the substitution of one non-polar (hydrophobic) residue such as alanine, isoleucine, valine, leucine or methionine for another, the substitution of one polar (hydrophilic) residue for another such as between arginine and lysine, between glutamine and asparagine, between glutamine and glutamic acid, between asparagine and aspartic acid, between glycine and serine, the substitution of one basic residue such as lysine, arginine or histidine for another, or the substitution of one acidic residue, such as aspartic acid or glutamic acid for another. The phrase "conservative substitution" also includes the use of a chemically derivatized residue in place of a non-derivatized residue provided that such polypeptide displays the requisite activity.

The term "derivative" refers to a peptide having one or more residues chemically derivatized by reaction of a functional side group. Such derivatized molecules include for example, those molecules in which free amino groups have been derivatized to form amine hydrochlorides, p-toluene sulfonyl groups, carbobenzoxy groups, t-butyloxycarbonyl groups, chloroacetyl groups or formyl groups. Free carboxyl groups may be derivatized to form salts, methyl and ethyl esters or other types of esters or hydrazides. Free hydroxyl groups may be derivatized to form O-acyl or O-alkyl derivatives. The imidazole nitrogen of histidine may be derivatized to form N-im-benzylhistidine. Also included as derivatives are those peptides which contain one or more naturally occurring amino acid derivatives of the twenty standard amino acids. For examples: 4-hydroxyproline may be substituted for proline; 5-hydroxylysine may be substituted for lysine; 3-methylhistidine may be

substituted for histidine; homoserine may be substituted for serine; and ornithine may be substituted for lysine.

MUC1 ligands of the present invention also include any peptide having one or more additions and/or deletions or residues relative to the sequence of a peptide whose sequence is shown herein, so long as the requisite activity is maintained or increased.

The term "fragment" refers to any subject peptide having an amino acid residue sequence shorter than that of a peptide whose amino acid residue sequence is shown herein.

By studying the binding sites recognized by the MUC1 ligands of the invention, the inventors determined that the ligands likely bind dipeptide motifs Ala-Pro (SEQ.ID.NO.: 28) and/or Pro-Ala (SEQ.ID.NO.:29) within the MUC1 tandem repeat. Specifically, MUC1-Lp1 binds the motif Pro-Ala-X-Gly-X-Thr (SEQ.ID.NO.:30) wherein X can be any amino acid.

Accordingly, the present invention provides a MUC1 ligand which binds to the sequence Ala-Pro or Pro-Ala. In a preferred embodiment, the MUC1 ligand binds to the sequence Pro-Ala-X-Gly-X-Thr, wherein X can be any amino acid.

2. Method of Isolating Ligands

As mentioned previously, the present inventors have isolated several MUC1 ligands using an improved screening assay. Importantly, using a traditional phage display technique wherein the MUC1 target is biotinylated and applied to streptavidin coated plates and then probed with a phage display library, no peptides were isolated. The improvement in the present method is due to the target used and the assay conditions employed.

As a target, the inventors used several repeats of the MUC tandem repeat, PDTRPAPGSTAPPAHGVTSA (SEQ.ID.NO.:1), coated directly on plates. Such a target is improved over traditional biotinylated MUC1 bound to streptavidin as it allows MUC1 to adopt a proper secondary structure. In particular, a MUC1 peptide comprising three consecutive MUC1 repeats adopts a secondary structure composed of

repeating knob-like structures (type IIβ turns) connected by extended spacers. The knobs protrude away from the long axis of the mucin with the most predominant antigenic site, APDTR (SEQ.ID.NO.:31), forming the accessible tip of the knob. As a result, such a target represents a constrained, well defined and attractive target to search for ligands. The MUC1 tandem repeat target may be glycosylated, partially glycosylated or unglycosylated.

The constrained target was used to probe a phage display peptide library. Several rounds of phage panning were done, each at higher salt concentrations. The phage that bound the MUC1 target at 400mM salt concentration were isolated. As a result the method isolated ligands that bind to the MUC1 target under high stringency conditions.

Accordingly, the present invention provides a method of isolating a ligand that binds to MUC1 comprising:

(a) providing a MUC1 target molecule wherein the MUC1 target molecule comprises at least 3 tandem repeats of the sequence PDTRPAPGSTAPPAHGVTSA (SEQ.ID.NO.:1);

(b) contacting the MUC1 target molecule with a peptide or small molecule library, under conditions to allow the MUC1 target to bind to any MUC1 ligands in the library, to prepare a test sample;

(c) panning the test sample under high stringency conditions; and

(d) isolating the MUC1 ligands bound to the MUC1 in the test sample.

The library is preferably a phage display peptide library but may also be a combinatorial peptide library or a combinatorial small molecule library.

The high stringency conditions preferably consist of a round of panning at high salt concentrations, for example about 200-400mM NaCl, preferably at least 300 mM NaCl. It will be appreciated by one of skill in the art that high stringency conditions can consist of any condition that can

destabilize protein interactions such as altering pH or using detergents, organic solvents or reducing agents.

Generally in step (c) the sample is panned in several rounds of panning, with the stringency conditions for phage binding increased at each round. Preferably there are 3 rounds of panning, each of reduced incubation time. In the first round of panning, phage binding is preferably performed at 75 mM NaCl; in the second round, the phage binding is preferably performed at 150 mM NaCl; and in the third round, the phage binding is preferably performed at least 300 mM NaCl. More preferably there are 4 rounds of phage panning wherein a second round is conducted at the same biochemical conditions as the first round although at a decreased incubation time. A preferred method of the phage panning strategy of the present invention is shown in Table 3.

The present invention also includes any ligand isolated according to the method including, and in addition to, the ligands described under 1. Ligands of the Invention.

Once potential ligands have been isolated, screening methods may be designed in order to determine if the molecules that bind to the MUC1 and are useful in the methods of the present invention. For example, the ligands can be tested to see if they bind MUC1 or MUC1 expressing cells as described in the Examples herein.

3. Uses of the Ligands

The present invention includes all uses of the ligands that bind to MUC1 as described herein. Some of the uses include the diagnosis, prognosis and treatment of cancer as well as the preparation of antibodies.

(a) Diagnosis or Monitoring of Cancer

In one embodiment, the present invention provides a use of a ligand to MUC1 to diagnose a cancer that is associated with MUC1.

The term "a cancer that is associated with MUC1" means any type of cancer wherein the cancer or tumour cells have underglycosylated MUC1 on their surface and/or they secrete or shed MUC1 from their

surface. Such cancers include breast, lung, colorectal, gastric, liver and pancreatic carcinomas.

In another embodiment, the present invention provides a use of a ligand to MUC1 as a prognostic tool or to monitor the efficacy of treatment of a cancer that is associated with MUC1.

The present invention also provides a method of diagnosing or monitoring a cancer that is associated with MUC1 comprising contacting a sample from a patient with a MUC1 ligand of the invention and assaying for binding between the MUC1 ligand and MUC1 in the sample, if present, wherein the presence of MUC1 indicates the presence of a cancer that is associated with MUC1. Samples which may be tested include blood, urine, serum, tears, saliva, feces, tissues and the like. For solid tumors a tissue biopsy may be used.

The method used to detect the cancer can be any method known in the art. For example, the binding of the ligands with MUC1 may be used in known immunoassays which rely on the binding interaction between a ligand of the invention and MUC1. Examples of such assays are radioimmunoassays, enzyme immunoassays (e.g. ELISA), immunofluorescence, immunoprecipitation, latex agglutination, hemagglutination, and histochemical tests. The resulting MUC1 bound to the ligand may be isolated by conventional isolation techniques, for example, salting out, chromatography, electrophoresis, gel filtration, fractionation, absorption, polyacrylamide gel electrophoresis, agglutination, or combinations thereof. Thus, the ligands may be used to identify or quantify the amount of a MUC1 in a sample in order to diagnose the presence of a tumor.

For use in diagnostic assays, the ligand may be labelled with a detectable marker including various enzymes, fluorescent materials, luminescent materials and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, biotin, alkaline phosphatase, ß-galactosidase, or acetylcholinesterase; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate,

rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; and examples of suitable radioactive material include S-35, Cu-64, Ga-67, Zr-89, Ru-97, Tc-99m, Rh-105, Pd-109, In-111, I-123, I-125, I131, Re-186, Au-198, Au-199, Pb-203, At-211, Pb-212 and Bi-212. The ligands may also be labelled or conjugated to one partner of a ligand binding pair. Representative examples include avidin-biotin and riboflavin-riboflavin binding protein. Methods for conjugating or labelling the ligands discussed above with the representative labels set forth above may be readily accomplished using conventional techniques.

The ligands of the invention may be used in immuno-histochemical analyses, for example, at the cellular and sub-subcellular level, to detect MUC1, to localise it to particular cells and tissues, and to specific subcellular locations, and to quantitate the level of expression. The ligands of the invention may also be used in flow cytometry.

Cytochemical techniques known in the art for localizing antigens using light and electron microscopy may be used to detect MUC1. Generally, a ligand of the invention may be labelled with a detectable substance and MUC1 may be localised in tissues and cells based upon the presence of the detectable substance. Examples of detectable substances include various enzymes such as biotin, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; fluorescent materials such as fluorescein; luminescent materials such as luminol; and, radioactive materials such as radioactive iodine I^{125}, I^{131} or tritium. Ligands may also be coupled to electron dense substances, such as ferritin or colloidal gold, which are readily visualised by electron microscopy.

The ligand may be immobilized on a carrier or solid support such as nitrocellulose, glass, polyacrylamides, gabbros, and magnetite. The support material may have any possible configuration including spherical (e.g. bead), cylindrical (e.g. inside surface of a test tube or well, or the external surface of a rod), or flat (e.g. sheet, test strip). When an insolubilized ligand is used protein bound to the ligand is isolated by

washing. For example, when the sample is blotted onto a nitrocellulose membrane, the ligand bound to MUC1 is separated from the unreacted ligand by washing with a buffer, for example, phosphate buffered saline (PBS) with bovine serum albumin (BSA).

Indirect methods may also be employed in which the ligand - MUC1 reaction is amplified by the introduction of an antibody, having specificity for the ligand.

When labelled ligand is used, the presence of a tumor can be determined by measuring the amount of labelled ligand bound to MUC1 in the sample or of the unreacted labelled ligand. The appropriate method of measuring the labelled material is dependent upon the labelling agent.

Where a radioactive label is used as a detectable substance, a MUC1 may be localized by radioautography. The results of radioautography may be quantitated by determining the density of particles in the radioautographs by various optical methods, or by counting the grains.

When unlabelled ligand is used in the method of the invention, the presence of MUC1 can be determined by measuring the amount of ligand bound to MUC1 using substances that interact specifically with the ligand to cause agglutination or precipitation. In particular, labelled antibody against a ligand of the invention, can be added to the reaction mixture. The presence of MUC1 can be determined by a suitable method from among the already described techniques depending on the type of labelling agent.

The reagents suitable for carrying out the diagnostic methods of the invention may be packaged into convenient kits providing the necessary materials, packaged into suitable containers. Such kits may include all the reagents required to detect MUC1 in a sample by means of the methods described herein, and optionally suitable supports useful in performing the methods of the invention.

(b) Therapeutic Uses

In a further embodiment, the present invention provides a use of a ligand to MUC1 to treat or prevent a cancer that has increased MUC1 levels. In such an embodiment the ligand may be coupled to an anticancer agent and used to target the agent to the cancer.

Accordingly, the present invention provides a use of a MUC1 ligand to prepare a medicament to treat or prevent a cancer that is associated with MUC1. The present invention also provides a use of a MUC1 ligand to treat or prevent a cancer that is associated with MUC1. The invention further includes a method of treating or preventing a cancer associated with MUC1 comprising administering an effective amount of a MUC1 ligand coupled to an anticancer agent to an animal in need thereof.

The term "animal" as used herein includes all members of the animal kingdom. Preferably, the animal is a mammal, more preferably a human.

The ligands coupled to anticancer agents may be formulated into pharmaceutical compositions for adminstration to subjects in a biologically compatible form suitable for administration *in vivo*. By "biologically compatible form suitable for administration *in vivo*" is meant a form of the substance to be administered in which any toxic effects are outweighed by the therapeutic effects. The substances may be administered to living organisms including humans, and animals. Administration of an effective amount of the pharmaceutical compositions of the present invention is defined as an amount effective, at dosages and for periods of time necessary to achieve the desired result. For example, a therapeutically active amount of a substance may vary according to factors such as the disease state, age, sex, and weight of the individual, and the ability of antibody to elicit a desired response in the individual. Dosage regima may be adjusted to provide the optimum therapeutic response. For example, several divided doses may be

administered daily or the dose may be proportionally reduced as indicated by the exigencies of the therapeutic situation.

The active substance may be administered in a convenient manner such as by injection (subcutaneous, intravenous, etc.), oral administration, inhalation, transdermal application, or rectal administration. Depending on the route of administration, the active substance may be coated in a material to protect the compound from the action of enzymes, acids and other natural conditions which may inactivate the compound.

The compositions described herein can be prepared by *per se* known methods for the preparation of pharmaceutically acceptable compositions which can be administered to subjects, such that an effective quantity of the active substance is combined in a mixture with a pharmaceutically acceptable vehicle. Suitable vehicles are described, for example, in Remington's Pharmaceutical Sciences (Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, Pa., USA 1985). On this basis, the compositions include, albeit not exclusively, solutions of the substances in association with one or more pharmaceutically acceptable vehicles or diluents, and contained in buffered solutions with a suitable pH and iso-osmotic with the physiological fluids.

(c) **Antibodies**

The ligands of the invention may also be used to prepare antibodies to the ligands. Conventional methods can be used to prepare the antibodies including polyclonal antisera or monoclonal antibodies. To produce polyclonal antibodies, a mammal, (e.g., a mouse, hamster, or rabbit) can be immunized with an immunogenic form of the ligand which elicits an antibody response in the mammal. Techniques for conferring immunogenicity on a peptide include conjugation to carriers or other techniques well known in the art. For example, the peptide can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum.

Standard ELISA or other immunoassay procedures can be used with the immunogen as antigen to assess the levels of antibodies. Following immunization, antisera can be obtained and, if desired, polyclonal antibodies isolated from the sera.

To produce monoclonal antibodies, antibody producing cells (lymphocytes) can be harvested from an immunized animal and fused with myeloma cells by standard somatic cell fusion procedures thus immortalizing these cells and yielding hybridoma cells. Such techniques are well known in the art, (e.g., the hybridoma technique originally developed by Kohler and Milstein (Nature 256, 495-497 (1975)) as well as other techniques such as the human B-cell hybridoma technique (Kozbor et al., Immunol. Today 4, 72 (1983)), the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., Monoclonal Antibodies in Cancer Therapy (1985) Allen R. Bliss, Inc., pages 77-96), and screening of combinatorial antibody libraries (Huse et al., Science 246, 1275 (1989)). Hybridoma cells can be screened immunochemically for production of antibodies specifically reactive with the peptide and the monoclonal antibodies can be isolated. Therefore, the invention also contemplates hybridoma cells secreting monoclonal antibodies with specificity for a protein of the invention.

The term "antibody" as used herein is intended to include fragments thereof which also specifically react with a protein, of the invention, or peptide thereof. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above. For example, $F(ab')_2$ fragments can be generated by treating antibody with pepsin. The resulting $F(ab')_2$ fragment can be treated to reduce disulfide bridges to produce Fab' fragments.

Chimeric antibody derivatives, i.e., antibody molecules that combine a non-human animal variable region and a human constant region are also contemplated within the scope of the invention. Chimeric antibody molecules can include, for example, the antigen binding domain from an antibody of a mouse, rat, or other species, with human constant

regions. Conventional methods may be used to make chimeric antibodies containing the immunoglobulin variable region which recognizes a protein of the invention (See, for example, Morrison et al., Proc. Natl Acad. Sci. U.S.A. 81,6851 (1985); Takeda et al., Nature 314, 452 (1985), Cabilly et al., U.S. Patent No. 4,816,567; Boss et al., U.S. Patent No. 4,816,397; Tanaguchi et al., European Patent Publication EP171496; European Patent Publication 0173494, United Kingdom patent GB 2177096B).

Monoclonal or chimeric antibodies specifically reactive with a ligand of the invention as described herein can be further humanized by producing human constant region chimeras, in which parts of the variable regions, particularly the conserved framework regions of the antigen-binding domain, are of human origin and only the hypervariable regions are of non-human origin. Such immunoglobulin molecules may be made by techniques known in the art, (e.g., Teng et al., Proc. Natl. Acad. Sci. U.S.A., 80, 7308-7312 (1983); Kozbor et al., Immunology Today, 4, 7279 (1983); Olsson et al., Meth. Enzymol., 92, 3-16 (1982)), and PCT Publication WO92/06193 or EP 0239400). Humanized antibodies can also be commercially produced (Scotgen Limited, 2 Holly Road, Twickenham, Middlesex, Great Britain).

Specific antibodies, or antibody fragments, reactive against a ligand of the invention may also be generated by screening expression libraries encoding immunoglobulin genes, or portions thereof, expressed in bacteria with the ligands of the invention. For example, complete Fab fragments, VH regions and FV regions can be expressed in bacteria using phage expression libraries (See for example Ward et al., Nature 341, 544-546: (1989); Huse et al., Science 246, 1275-1281 (1989); and McCafferty et al. Nature 348, 552-554 (1990)).

The following non-limiting examples are illustrative of the present invention:

EXAMPLES

Example 1

Preparation of MUC1 Ligands

Cell Lines

The human breast cancer cell line T47D (American Type Culture Collection; Rockville, MD) was cultured at 37°C in α-MEM, Iscove's modification of DMEM (IMDM), or RPMI 1640 in the presence of 5% CO_2 with 10% FCS (Cansera International Inc., Rexdale, Ontario). The expression of the MUC1 epitope on these cell lines was confirmed by flow cytometry using a series of anti-MUC1 mAb BC-2 (28,29).

Peptide Synthesis

All peptides were assembled by solid-phase peptide synthesis on an Applied Biosystems 341A Peptide Synthesizer using Wang resin supports (4-hydroxymethylphenoxymethyl copolystyrene -1% divinylbenzene; Applied Biosystems) and 9-fluorenyloxycarbonyl protected (Fmoc) amino acids (NovaBiochem, San Diego, CA). Two MUC1 peptides were constructed, namely a 60-amino acid (MUC1-60mer) and a 100-amino acid long peptide (MUC1-100mer) corresponding to three and five consecutive tandem repeats of the MUC1 mucin. Each repeat was composed of the sequence Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala. The peptides SWWPFPPQPDPA (SEQ.ID.NO.:3) and YSWWPFPPQPDPA (SEQ.ID.NO.:41) corresponding to sequences derived from our phage display searches were synthesized by Fmoc chemistry using 2-chloro-trityl chloride resin (200-400 mesh; NovaBiochem).

Some peptides were biotinylated at their N-terminus by adding two N-terminal glycines to the sequence of these peptides and by subsequently treating the deprotected peptide resins (0.125 mmoles) with 3 molar equivalents of biotin (92 mg) and HATU/HOBt (1:1 ratio; 200 mg) dissolved in 5 mL N-methylpyrrolidone. After 5 minutes, 6 molar equivalents of diisopropylethylamine were added and the reaction left to occur overnight with constant stirring. Peptides were released from their

supports by suspending and mildly agitating each peptide resin in 82.5% (v/v) TFA: 5% (v/v) phenol: 5% (v/v) thioanisole: 5% (v/v) water:2.5% (v/v) EDT for 2 hours at room temperature. Each treated resin was filtered and the cleaved peptide was extracted from the resin with 1 mL TFA. The subsequent filtrate was recovered and the peptide precipitated with diethylether. Crude peptides were redissolved in water and lyophilized. The MUC1 peptides were purified to homogeneity by reverse phase HPLC (Waters HPLC 510 System). Analytical separations were performed on a Rapid Resolution Zorbax SB-C_{18} column (4.6 x 75mm; 3.5 µ particles) at a flow rate of 1 mL/min while semi-preparative separations were carried out on a Zorbax SB-C_{18} column (9.4 x 250mm; 5µ particles) operating at a flow rate of 2 mL/min. In the case of MUC1-100mer, the columns were equilibrated in solvent A (0.1% TFA/AcN) and the peptide was eluted using a 30-minute gradient going from 10% to 35 % mobile phase B (0.1%TFA/60%AcN). For the MUC1-60mer, the columns were equilibrated in 0.1% TFA/20%AcN (eluent A) and the peptide was resolved using a 28-minute gradient going from 0% to 28 % eluent B (0.1% TFA/AcN). For all other peptides, the same columns were equilibrated in 95% eluent A: 5% eluent B and the peptides were eluted using a 40-minute gradient going from 5% to 60% eluent B. The mass and composition of MUC1 peptides were confirmed by mass spectrometry and amino acid analysis. MUC1-60mer, MH+ observed 5625.1 (calc. 5625.1), biotinylated homolog, MH+ observed 5851.9 (calc. 5851.4); MUC1-100mer, MH+ observed 9362 (calc.9363). Amino acid composition: MUC1-60mer; Asp, found 0.8 (exp. 1); Ser, found 2.1 (exp. 2); Gly, found 2.1 (exp. 2); Arg, found 1.0 (exp.1); Thr, found 2.6 (exp.3); Ala, found 4.2 (exp. 4); Pro, found 4.8 (exp. 5); Val, found 1.0 (exp. 1); His, N.D; MUC1-100mer; Asp, found 1.1 (exp. 1); Ser, found 2.4 (exp. 2); Gly, found 2.4 (exp. 2); Arg/Thr, found 4.0 (exp.4); Ala, found 3.7 (exp. 4); Pro, found 4.4 (exp. 5); Val, found 1.1 (exp. 1); His, N.D. The mass of other peptides were also confirmed by mass spectrometry: TMGFTAPRFPHY, MH+ observed 1424.2 (calc.1424.6); biotinylated homolog, MH+ observed 1765.5 (calc. 1765.7);

YTKSDPLKLLES, MH+ observed 1393.4 (calc.1393.5); biotinylated homolog, MH+ observed 1733.4 (calc. 1734.6); SWWPFPPQPDPA, MH+ observed 1424.2 (calc. 1424.6); biotinylated homolog, MH+ observed 1765.7 (calc. 1765.7); YSWWPFPPQPDPA, MH+ observed 1586.8 (calc.1587.7).

Pin Synthesis

Overlapping octapeptides spanning the entire sequence of the 20-amino acid long MUC1 tandem repeat were synthesized by Fmoc chemistry on derivatized high density polyethylene pins (Epitope Scanning kit; Cambridge Research Biochemicals, Wilmington, DE) following methods described previously (13-15). Six copies of each peptide were prepared for statistical reasons (see section on ELISA results). Pins harboring the peptide sequences PLAQ (positive control using an antipeptide antibody provided by the supplier) and GLAQ (negative control) were synthesized as control sequences to monitor the quality of the synthesis.

Panning strategy of phage display library

Phage library searches were performed using the Ph.D-12 Peptide Library (New England Biolabs, Beverly, MA), a M13 bacteriophage library displaying random 12-amino acid long linear peptides fused to each phage coat protein. The protocols used in phage panning followed guidelines suggested by the manufacturer. Briefly, polyvinyl chloride 96-well plates were coated with the 100-amino acid long synthetic MUC1 (5 repeats of the core sequence) in 0.1 M carbonate buffer, pH 9.6, at 37°C for 30 minutes or 4°C overnight. Wells were subsequently treated for 30 minutes at 37°C with PBS containing 0.5% (w/v) bovine serum albumin (BSA) and 0.05% (v/v) Tween-20. Plates were then washed with PBS and 100 µL aliquot of the phage library-containing solution were dispensed in each well ($2x10^{11}$ pfu/ml binding buffer). The stringency condition used for phage binding was altered at each round of panning. During the first round of panning, phage binding was performed overnight at room temperature in 10 mM PBS, pH 7.4 containing 75 mM sodium chloride, 0.5% (w/v) bovine serum albumin (BSA) and 0.05% Tween-20. The

second round of phage binding was shortened to 2 hours using the same conditions. The third round of panning binding was performed at RT for 30 minutes in PBS, pH 7.4, containing 0.5% (w/v) BSA, 0.05% (V/V) Tween-20. The last round of panning was done at room temperature for 30 minutes in 10mM PB pH7.4 containing different concentrations of NaCl (200mM, 250mM, 300mM, 350mM, 400mM). Between each panning round, the plates were washed 10 times at room temperature in the same binding buffer as used in the panning assay. The bound phages were eluted by dispensing 100 µl of 0.2M Glycine-HCl,pH 2.2, containing 1% BSA in each well. One microliter of the phage eluate was used to titer the phage and the remaining eluate was amplified. After the last round of panning, clones recovered from the highest binding stringency conditions were picked and amplified. Phage DNA were prepared and sequenced.

Mapping of peptides and phage binding sites on the MUC1 tandem repeat

MUC1 peptides covalently attached to polyethylene pins by their C-terminus were incubated in blocking buffer (1% [wt/vol] bovine serum albumin, 1% [wt/vol] ovalbumin, 0.1% [vol/vol] Tween 20 in 10 mM phosphate-buffered saline [PBS, pH 7.4]) for 1 hour to prevent the non specific absorption of phages, peptides or secondary antibodies. The pins were subsequently incubated overnight at 4°C in wells containing 150 µL aliquots of solutions of either purified phages or related biotinylated synthetic peptides (µg/mL) prepared in blocking buffer. After 4 washes in PBS containing 0.05% (vol/vol) Tween 20, the pins were incubated for 1 hour in wells containing 100 µL of either rabbit anti-phage antibody (for detecting phages, 1:1000 dilution) or streptavidin-HRP conjugate (for detecting peptides, 1:1000 dilution). In the case of phage detection, the pins were further washed and incubated for 1 hour with a goat anti-rabbit immunoglobulin-peroxidase construct (100 µL of a 1:1000 dilution) diluted in blocking buffer. All pins were finally washed and the binding of phages or peptides was detected by incubating the pins in microtiter plates containing 100 µL of 0.05% (wt/vol) 2,2′-azino-bis(3-ethylbenzthiazoline-6-sulfonate) (ABTS) dissolved in 0.1 M sodium phosphate-0.08 M citric acid

(pH 4.0) - 0.003% (vol/vol) hydrogen peroxide per well. Absorbance readings at 405 nm of solutions in each well were recorded with a Titertek plate reader. Each point represents the average absorbance reading calculated for assays performed on six copies of each overlapping MUC1 octapeptide.

Flow Cytometry

Breast cancer cells (1-2 x 10^5 cells; T47D) were suspended in 50 µl of PBS (pH 7.4; supplemented with 1mM $CaCl_2$, 1mM $MgCl_2$, 1% (w/v) BSA, 0.05% sodium azide) containing either 200 µM of a biotinylated peptide or the monoclonal antibody BC-2 (IgG1, 1:100 dilution; Serotec, Oxford, England). The cells were incubated at 37°C for 60 minutes, subsequently washed three times with PBS, resuspended in 50 µl of the same buffer containing 1 µL of 1 mg/mL phycoerythrin (PE)-streptavidin conjugate (Molecular Probes; Eugene, OR) or 1 µL of a 0.2 mg/mL PE-antimouse Ig kappa light chain conjugate (PharMingen, LaJolla, CA) and finally incubated at 4°C for 30 minutes. The cells were washed three times with PBS and resuspended in 200 µl of PBS containing 5 µl of 7-aminoactinomycin D (7-AAD, 0.1 mg/mL in PBS, pH 7.4). Flow cytometry was performed on a Becton Dickinson FACScan and the data analysed using CellQuest software.

Radioligand Binding Assays

Synthetic, tyrosine-containing MUC1 peptide ligands were iodinated using the chloramine T method (32). Briefly, 100 µg of peptide dissolved in 20 µL of water was dispensed into polypropylene tube. A 10 µL aliquot of $Na^{125}I$ (1mCi; Dupont) as well as 10 µL of chloramine T (BDH; 2mg/mL in water) were added to the tube and the iodination reaction was allowed to proceed for 1 minute at room temperature in a fumehood. The reaction was stopped with 10 µL of sodium metabisulfite (4 mg/mL in water) and the mixture diluted to 1 mL with 0.1% (v/v) TFA/water. The radiolabeled peptides were loaded and desalted on disposable C18 cartridges (Millipore) in the same eluent. The labeled peptides were then eluted from the cartridges with 100% methanol and

concentrated by evaporating the organic phase under a stream of nitrogen gas. The purity of labeled peptides was > 95% as verified by thin layer chromatography and their specific activity range in value from 15-20 μCi/μmol. The breast cancer cell line T47D expresses MUC1 on its surface and was used for ligand binding assays and competition assays. For binding assays, increasing concentrations of each radiolabeled peptide in the presence or absence of an excess unlabeled peptide (200 fold excess) were incubated with 10^5 cells in 0.2 ml of binding buffer (PBS, pH7.4, 1mM $CaCl_2$, 1mM $MgCl_2$, 1% (w/v) BSA, 0.05% sodium azide) at room temperature for 60 minutes. The reaction mixture were spun through an oil column (84 % [v/v] silicon oil and 16% paraffin oil). After freezing the column on dry ice for 20 minutes, the tip of the column containing the cell pellet was cut off. The radioactivity of both the tip and the remaining column (supernatant) were counted using a Gamma counter. Peptide binding constants were derived from titration curves using GRAPHPAD software. For competition assays, radioiodinated peptides were incubated with different concentrations of unlabeled peptides (MUC1 peptide ligands or the 100mer synthetic MUC1 peptide) and dispensed into tubes containing 5 x 10^4 T47D cells suspended in 150 μL of binding buffer. Displacement curves were constructed from experiment performed in duplicate.

Plate Binding Assays

Wells of 96-well polystyrene plates (NUNC) were coated with 100 μL of streptavidin (SIGMA; 10 μg/ml solution in 0.1 M carbonate buffer [pH 9.6]), at either 37°C for 30 minutes or overnight at 4°C. The plates were washed three times with PBS, and the biotinylated 60-mer MUC1 peptide (0.5 μg/mL; 100 μL) prepared in binding buffer, was added to the wells and incubated at 37°C for 30 minutes. The mAb BC-2 (100 μL; 1:5000 dilution) pre-mixed with increasing concentrations of a MUC1 peptide ligand (1 μM - 1000μM) were dispensed in MUC1-coated wells and the binding event left to proceed at 37°C for 60 minutes. The plates were subsequently washed with PBS. Aliquots of anti-mouse IgG-HRP

conjugate (100 µL; 1:500 dilution; SIGMA) were placed in each well and incubated for 30 minutes at room temperature. Following a final washing step in PBS, 100 µL aliquot of the color substrate 3,3',5,5'-tetramethylbenzidine (TMB, 100 µg/ml in 0.1M sodium acetate buffer, pH5.5, 1.3mM H_2O_2) was placed in each well to reveal the presence of the peroxidase enzyme. The reaction was stopped after a 30-minute incubation period at room temperature by adding 100 µL of a 2M H_2SO_4 solution. Absorbance readings were recorded at 450 nm using a plate reader (Titertek Multiskan MCC/340).

Results of the Screening Strategy

The inventors have synthesized by solid phase peptide synthesis and purified a 100 amino acid-long analog of MUC1 which contains five consecutive tandem repeats ($[PDTRPAPGSTAPPAHGVTSA]_5$). The objective of this study was to identify one or more peptide ligands able to recognize and bind specifically to the mucin MUC1 tandem repeat. One solution was to perform searches within phage display library for potential peptide ligands. The success in finding such ligands depends in part on the structure and nature of the target itself. The 100-amino acid long peptide representing 5 consecutive MUC1 tandem repeats (MUC1 100mer) was thus synthesized in order to ensure the presence of secondary structure observed in the MUC1 mucin. A phage library was selected where random 12-amino acid sequences are presented as fused domains onto the M13 coat protein. Panning procedures were devised where reduced incubation times and increasing salt concentrations were used during phage binding steps prior to the elution of bound phages.

The screening of a phage display peptide library for sequences for possible ligands to the MUC1 peptide tandem repeat identifies several distinct 12-amino acid long peptide sequences which are shown in Table 1. Three distinct peptide sequences derived from screening a phage display peptide library were shown to bind to the MUC1 peptide tandem repeat. These are identified as MUC1-Lp1; MUC1-Lp2 and MUC1-Lp3 in Table 1.

The binding of these peptides to MUC1 expressed on the breast cancer cell line T47D was shown to be specific (displaced with the corresponding unlabeled analogs) with affinity constants in the micromolar range.

Based on the sequence of the ligands isolated, the inventors identified a motif that was common to the ligands. The motif in each ligand is shown in Table 2. Generally, the motif comprises the sequence X_1-X_2-X_3 wherein X_1 is Pro, Thr or Ser; X_2 is any amino acid; and X_3 is Pro, Ala, Lys, His, Thr, Asn, Tyr or Phe. Preferably, X_1 is Pro; X_2 is a hydrophobic amino acid (i.e. Ala, Val, Leu, Ile, Pro, Met, Phe or Try) and X_3 is Pro.

Discussion

The underglycosylated or deglycosylated forms of the MUC1 tandem repeat represent targets for immunotherapy or for designing cancer vaccines. The value of the MUC1 repeat as an immunogen for the generation of useful B- or T-cell responses is presently being evaluated in the context of clinical trials. More than fifty monoclonal antibodies have also been raised against surface determinants on epithelial cell cancers including breast, colorectal and ovarian cancers, and been shown to recognize a common short peptide region of the MUC1 tandem repeat. From practical reasons, the use of these antibodies as therapeutic agents is hampered by their size, affinity and structure leading to issues such as their immunogenicity and their reduced ability to penetrate into tumors. Small ligands able to bind specifically to tumor-associated forms of the MUC1 epitope would be advantageous in these respects.

The challenge was to identify small peptides able to bind to the 20-amino acid tandem repeat of MUC1, a sequence that is predicted to adopt a minimal but potentially unique secondary structure. Using a phage display strategy, the inventors identified many peptides able to bind specifically to the determinant found on the MUC1 tandem repeat. These peptides offer the potential for developing diagnostic/therapeutic low molecular-weight ligands directed at MUC1 sites present on cancer cells.

The MUC1 tandem repeat represents one of the best defined target in terms of developing guided cancer therapies. Recent studies have suggested the need to identify low molecular agents that can target cells expressing underglycosylated forms of the MUC1 mucin.

Mucins serve a range of functions associated with the physical protection of mucosal surfaces (lubrication), in coping with rapid changes in pH and osmolarity, or in binding via its carbohydrate side chains to pathogenic strains of bacteria or viruses or in blocking the activity of neutrophils, cytotoxic T cells and NK cells.

Example 2

Defining the binding sites of three MUC1 ligands on the MUC1 tandem repeat

The binding of MUC1-Lp1 (SEQ.ID.NO.:2), MUC1-Lp2 (SEQ.ID.NO.:3), and MUC1-Lp3 (SEQ.ID.NO.:4) to the MUC1 tandem repeat was confirmed and further delineated using overlapping octapeptides assembled on plastic pins, covering the entire sequence of the 20-amino acid long MUC1 tandem repeat. Biotinylated analogs of these MUC1 ligands were synthesized and shown to bind to more than one region of the tandem repeat (Figure 1). Their binding patterns suggest that they associate with short, duplicated motifs present within the MUC1 sequence (Figure 1). More precisely, the two dominant epitopes recognized by MUC1-Lp1 were Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr (SEQ.ID.NO.:32) and Ala-Pro-Pro-Ala-His-Gly-Val-Thr (SEQ.ID.NO.:33), sequences which harbour the common motif Pro-Ala-X-Gly-X-Thr (SEQ.ID.NO.:30). In the case of MUC1-Lp2, the MUC1 peptides recognized were Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr (SEQ.ID.NO.:32), Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala (SEQ.ID.NO.:34), Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro (SEQ.ID.NO.:35) and Ala-His-Gly-Val-Thr-Ser-Ala-Pro (SEQ.ID.NO.:36) suggesting single or combinations of two possible binding units namely Ala-Pro (SEQ.ID.NO.:28) and Gly-X-Thr (SEQ.ID.NO.:37). MUC1-Lp3, unlike the previous two ligands, appears to bind to a broad array of MUC1 sequences containing Ala-Pro (SEQ.ID.NO.:28) or Pro-Ala (SEQ.ID.NO.:29) dipeptides.

In summary, epitope mapping studies indicate that all three synthetic ligands may recognize at least in part the dipeptides Ala-Pro and/or Pro-Ala. MUC1-Lp3 requires the presence of only these dipeptide motifs within the MUC1 repeat. Interestingly, most monoclonal antibodies that bind to the tumor-associated MUC1 repeat recognize parts or all of the sequence Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro (SEQ.ID.NO.:38) where Ala-Pro-Asp-Thr-Arg (SEQ.ID.NO.:39) represents an accessible bulge displayed on the surface of the underglycosylated MUC1 repeats (12). This region includes two Ala-Pro and one Pro-Ala sequences. The two MUC1 Gly-X-Thr regions defined in the MUC1-Lp1 and -Lp2 binding sites are located on both sides of the protruding knob-like structure of each repeat (27). Their location would suggest that searches for small MUC1 ligands may led to peptides associated with distinct and potentially less accessible regions of the MUC1 repeat.

Defining structural motifs explaining the association of MUC1 binders to the MUC1 tandem repeat

The specificity of our MUC1-ligand peptides (MUC1-Lp) for the MUC1 tandem repeat was also confirmed in a displacement assay where unlabelled MUC1-Lps were shown to inhibit the binding of a well known anti-MUC1 antibody to the synthetic MUC1/3TR peptide (Figure 2). Briefly, the mAb BC2 (28) recognizes the sequence Ala-Pro-Asp-Thr-Arg (SEQ.ID.NO.:39) within the MUC1 repeat, reacts with MUC1 components in the sera of patients with breast cancer, and binds avidly to synthetic analogs of the MUC1 tandem repeat with a reported affinity constant of 2.5 x 10^9 M^{-1} (29). Peptide mapping analyses (Figure 1) suggest that MUC1-Lp2 and MUC1-Lp3 peptides bind to the Ala-Pro dipeptide sequence present in the Ala-Pro-Asp-Thr-Arg epitope recognized by mAb BC2 (Figure 1) and as such these ligands should competitively inhibit the binding of mAb BC2 to MUC1/3TR. MUC1-Lp1 binds to sites in proximity to the mAb BC2 epitope and could also sterically interfere with this interaction. The relative inhibitory potential of individual MUC1 ligand peptides was measured in a displacement assay involving the binding of mAb BC2 to

MUC1/3TR. As expected, MUC1-Lp2 and Lp3 were better able to displace mAb BC2 bound to MUC1/3TR than MUC1-Lp1 while an irrelevant peptide (YTVTNSWTWWSPLQQA (SEQ.ID.NO.:40)) was unable to inhibit this interaction (Figure 2). Sets of three overlapping hexapeptides covering each of the MUC1-Lp sequences were synthesized to further define relevant residues present in MUC1-binding peptides. For MUC1-Lp1, the sequences Phe-Thr-Ala-Pro-Arg-Phe (SEQ.ID.NO.:17) and Pro-Arg-Phe-Pro-His-Tyr (SEQ.ID.NO.:18) were equal to MUC1-Lp1 as competitive inhibitors, suggesting that the common tripeptide Pro-Arg-Phe (SEQ.ID.NO.:19) may represent part of its putative MUC1-binding motif (Figure 2A). In the case of MUC1-Lp2, the peptide Ser-Trp-Trp-Pro-Phe-Pro (SEQ.ID.NO.:20) was comparable to the 12-amino acid ligand itself in competing with mAb BC2 for binding to the MUC1 repeat. The other two hexapeptides were weaker inhibitors and contained one or more tripeptides with the motif Pro-X-Pro (SEQ.ID.NO.:21) (Pro-Gln-Pro (SEQ.ID.NO.:22), Pro-Asp-Pro (SEQ.ID.NO.:23), Pro-Phe-Pro (SEQ.ID.NO.:24); Figure 2B). Finally, MUC1-Lp3 and to a lesser extent, Ser-Asp-Pro-Leu-Lys-Leu (SEQ.ID.NO.:25) and Tyr-Thr-Lys-Ser-Asp-Pro (SEQ.ID.NO.:26) were inhibitors of mAb BC2 association with MUC1 (Figure 2C). The recurrent tripeptide present in these MUC1-Lp3 analogs is the sequence Ser-Asp-Pro (SEQ.ID.NO.:27).

In summary, the recognition of Ala-Pro elements within the MUC1 repeat by all three peptide ligands would suggest the occurrence of a common structural motif within the sequences of MUC1-Lp1, -Lp2 and-Lp3 (Table 1) accounting for at least part of their MUC1-binding character. An analysis of antibody displacement curves by all three MUC1 binders and their respective hexapeptide analogs (Figure 2) revealed a tripeptide motif within MUC1-Lp1, -Lp2, and -Lp3 defined by either Pro or Ser at the first position followed by a charged/polar amino acid and ending with either a Pro or a Phe residue. All deduced insert sequences of MUC1-binding phages (Table 1) were then aligned to further defined the putative MUC1-ligand pattern. Several other sequences harboured motifs similar

to the ones initially deduced from the three most common sequences (Table 2). This analysis only led to a broadening of possible residues within these motifs, particularly at the second and third positions of the tripeptide segment. Nevertheless, the motif typically contains one or more proline residues preferentially located at the N- and/or C-terminus of the segment.

A search for proteins harboring the tripeptide motif interestingly identified the MUC1 tandem repeat itself as containing the related but not identical tripeptide motifs Pro-Asp-Thr, Pro-Ala-Pro, Thr-Ala-Pro, and Pro-Pro-Ala. This finding would suggest a possible homophilic aggregation of MUC1 molecules via their tandem repeats. MUC1/5TR however remains monomeric in solution as monitored by high performance gel filtration chromatography (results not shown). Other proline-rich sequences present in WW (Pro-X-Tyr) or SH3 (Pro-X-X-Pro) domain ligands partly resemble the MUC1 ligand motif (38,39). The motifs Pro-Pro-Gln and Pro-Gln-Pro are found in the tandem repeats of rhodopsin and gliadin respectively (40). The latter motifs are preceded by a Tyr or Phe residues and are part of an ensemble of proteins presenting a repetitive structural pattern defined as polyproline, ß turn helices (40). The interesting feature of this structural element appears to be its ability to exist as a series of local interconverting structures (ß-turn) centered at the Tyr/Phe-Pro bond and causing no major change in the overall polyproline helical character of the peptide chain. This built-in local flexibility may potentially represent a molecular switch associated with the recognition and posttranslational modification of tyrosine (phosphorylation), serine or threonine (glycosylation in mucin, phosphorylation) residues often found in such repeats. The tandem repeats located at the C-terminal of the RNA polymerase II as well as those of WW and SH3 domain ligands may adopt related structures in solution (41). In summary, the identified MUC1 ligand peptide motifs may possess an inherent potential to locally adapt to the conformation adopted by distinct peptide regions of the MUC1 tandem repeat.

Example 3

MUC1 ligands bind specifically to MUC1-expressing cells

MUC1-Lp1, -Lp2 and Lp3 were radio-iodinated to determine if these peptides could bind specifically to the MUC1 repeat presented on tumor cells. Labelling was facilitated by the presence of unique, naturally occurring C- or N-terminal tyrosine residues in the sequence of MUC1-Lp1 and -Lp3 respectively. A tyrosine was added to the N-terminus of the MUC1-Lp2 sequence to complete the analysis. Two cell lines were used to construct binding curves for all three labelled ligands: T47D, an established MUC1-expressing breast cancer cell line and Daudi, a human B-lymphoma cell line that does not express MUC1 mucin. Binding curves presented in Figure 3 highlight the fact that all three MUC1 ligands associated specifically with T47D cells while MUC1-negative Daudi cells did not express significant or detectable levels of ligand receptors. The specific binding of radiolabeled MUC1-Lps to T47D cells was saturable, as determined from the difference in counts (cpm) between total and non-specific binding. The dissociation constant and number of binding sites of each MUC1 ligands on T47D cells are presented in Table 4. The peptide ligands bind to T47D cells with K_d values in the micromolar range. MUC1-Lp1 represents the most avid MUC1 ligand with the fewest number of binding sites on T47D cells while MUC1-Lp2 and -Lp3 bind to significantly more sites on these cells suggesting that the motifs that they recognize on the MUC1 tandem repeat are less complex than the binding site for MUC1-Lp1. The association of MUC1 ligand peptides to T47D cells was reversible as defined by the concentration-dependent displacement of radiolabeled MUC1-Lp1, -Lp2 and -Lp3 from these cells with their respective unlabelled MUC1-Lps (Figure 4A). The distinct selectivity of MUC1-Lp1 for the MUC1 repeat was further confirmed by the fact that unlabelled MUC1-Lp1 was significantly better than MUC1-Lp2 and -Lp3 in competing with radio-iodinated MUC1-Lp1 for binding to T47D cells (Figure 4B).

The results presented in Table 4 provide a strong rationale for the sequence motifs deduced from our epitope mapping studies. More precisely, the dipeptide motifs Ala-Pro and Pro-Ala occur frequently within the MUC1 tandem repeat (2 Ala-Pro, 2 Pro-Ala, one additional Ala-Pro linking two repeats). MUC1-Lp3 was thus expected to bind to a higher number of sites on MUC1 than MUC1-Lp1 and MUC1-Lp2 which recognize additional residues. The association of MUC1-Lp3 to simple dipeptide motifs also implied that its affinity for the tandem repeat would be comparatively lower than MUC1-Lp1 and MUC1-Lp2. The tabulation of binding constants and number of binding sites for each ligand on MUC1-expressing breast cancer cells (T47D) support these predictions. MUC1-Lp3 was the weakest MUC1 binder and bound to 40 times more sites on MUC1-expressing cells than MUC1-Lp1 and to 3 fold more sites than MUC1-Lp2 confirming that it recognizes a more promiscuous target on the tandem repeat than the other two MUC1 ligands. As expected, MUC1-Lp1 which bound to a relatively defined site, has the fewest number of binding sites on T47D cells and displays a higher affinity and selectivity (Figure 4B) for MUC1 sites on cells than MUC1-Lp2 and -Lp3.

Prospects for deriving small ligands displaying high affinity and selectivity for the MUC1 tandem repeat

The frequent exposure of peptide domains in the MUC1 tandem repeat defines an important class of related tumor-associated epitopes on the surface of epithelial cancer cells. More importantly, the pattern of internalization and recycling of membranous MUC1 suggests that ligands targeting such sites will be routed inside MUC1-expressing cells and could serve as delivery vehicles (42). Single-chain Fv antibodies directed at the MUC1 core peptide tandem repeat have recently been identified (23). However, the mass of such structures remains high (> 10 kD). Issues relating to their ability to penetrate into solid tumors or to recognize sterically hindered peptide epitopes in the context of underglycosylated forms of MUC1-expressing tumor cells will need to be addressed. The design of peptide "mimics" of MUC1-directed antibodies

has recently been suggested to alleviate some of these issues (43). The MUC1 ligands identified in this study have affinity constants in the micromolar range suggesting that their MUC1 recognition motifs could be improved by constructing and probing phage display or synthetic peptides sub-libraries using MUC1-Lp sequences as starting templates. The selectivity of MUC1-Lp1 for the MUC1 tandem repeat appears to be based on a relatively complex discontinuous motif, namely Pro-Ala-X-Gly-X-Thr. This finding would suggest that an increase in affinity constant may also be gained by constraining the structure of MUC1 ligands through the use of peptide mimetics rather than through library searches involving a limited spectrum of amino acid building blocks.

In conclusion, short peptides that bind to the tumor-associated MUC1 tandem repeat were identified by screening a phage display peptide library Three synthetic peptides were shown to bind specifically to MUC1-expressing breast cancer cells but not to MUC1-negative cells. More importantly, the alignment of peptides derived from sequencing inserts of MUC1-binding phages as well as antibody inhibition assays suggest the existence of short proline-containing peptide motifs in the identified MUC1 binders. These peptides can now be used as lead templates in developing small ligands directed at MUC1 sites present on cancer cells.

While the present invention has been described with reference to what are presently considered to be the preferred examples, it is to be understood that the invention is not limited to the disclosed examples. To the contrary, the invention is intended to cover various modifications and equivalent arrangements included within the spirit and scope of the appended claims.

All publications, patents and patent applications are herein incorporated by reference in their entirety to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety.

TABLE 1

Duodecapeptide sequences deduced from sequencing phage DNA inserts of MUC1-binding phages

Peptide Sequence [1]	Frequency	SEQ.ID.NO.
TMGFTAPRFPHY (MUC1-Lp1)	12/34	2
SWWPFPPQPDPA (MUC1-Lp2	2/34	3
YTKSDPLKLLES (MUC1-Lp3)	8/34	4
VVPVHWSRGVVL	1/34	5
HIPVAALAPRMT	1/34	6
LGLQPPTSALDP	1/34	7
TPAFSPLPTDLL	1/34	8
ELNTHLATNVFT	1/34	9
IDVHSINFLATL	1/34	10
THPWSLKSTSFF	1/34	11
YITPYAHLAGGN	1/34	12
SLPIPSHARLQN	1/34	13
YLPYATLSQNSH	1/34	14
WHIPPNIGRTFS	1/34	15
TSNPHTRHYYPI	1/34	16

[1] One-letter amino acid code

TABLE 2

Possible tripeptide motifs that bind to the peptide core of the MUC1 tandem repeat

TPAFS	**PLP**	TDLL
SL	**PIP**	SHARLQN
SWW	**PFP**	PQPDPA
SWWPFPPQ	**PDP**	A
SWWPFP	**PQP**	DPA
SWWPF	**PPQ**	PDPA
LGLQ	**PPT**	SALDP
WHI	**PPN**	IGRTFS
YTKSD	**PLK**	LLES
VV	**PVH**	WSRGVVL
HI	**PVA**	ALAPRMT
YIT	**PYA**	HLAGGN
YL	**PYA**	TLSQNSH
TMGFTA	**PRF**	PHY
TSN	**PHT**	RHYYPI
TMGFTAPRF	**PHY**	
TMGF	**TAP**	RFPHY
	THP	WSLKSTSFF
YTK	**SDP**	LKLLES

TABLE 3

Phage panning strategy used to identify phages that bind specifically to the peptide core of the MUC1 tandem repeat

Round of Phage Panning	Conditions	Comments
1	- 10 mM phosphate buffer, 75 mM NaCl, pH 7.4, 0.5% (w/v) BSA, 0.05% Tween-20 - Overnight incubation at room temp. - Washed 10x with same buffer	- low stringency - high number of phage particles (few copies of each library element) - S/NS very low
2	- 10 mM phosphate buffer, 75 mM NaCl, pH 7.4, 0.5% (w/v) BSA, 0.05% Tween-20 - 2-hour incubation at room temp. - Washed 10x with same buffer	- higher stringency - low number of phage particles (higher number of copies of each residual library element) - S/NS much higher
3	- PBS, pH 7.4, 0.5% (w/v) BSA, 0.05% Tween-20 - 30-minute incubation at room temp. - Washed 10x with same buffer	- increased stringency - higher number of phage particles (enrichment for the best MUC1-binding elements of the library) - S/NS - 1 or greater
4	- 10 mM phosphate buffer, ≥ 300 mM NaCl, pH 7.4, 0.5% (w/v) BSA, 0.05% Tween-20 - 30-minute incubation at room temp. - Washed 10x with same buffer	- highest stringency - further enrichment for the best MUC1-binding elements of the library) - S/NS ›1

S/NS stands for the ratio of phages specifically bound to the target (S) versus non-specifically bound phages (NS).

TABLE 4

Binding parameters of peptide ligands to MUC1-expressing T47D cells

Peptide	K_d (μM)	Binding sites per cell (x 10^5)
MUC1-Lp1	1.9 (± 2.2)	1.5 (± 1.1)
MUC1-Lp2	2.4 (± 1.6)	22 (± 10)
MUC1-Lp3	8.3 (± 1.4)	60 (± 8.6)

FULL CITATIONS FOR REFERENCES REFERRED TO IN THE SPECIFICATION

1. Apostolopoulos, V., and McKenzie, I.F.C. (1994) *Crit. Rev. Immunol. 14,* 293-309.

2. Taylor-Papadimitriou, J. and Finn, O.J. (1997) *Immunol. Today 18,* 105-107.

3. Miles, D.W., and Taylor-Papadimitriou, J (1999) *Pharmacol. Ther. 82,* 97-106.

4. Taylor-Papadimitriou, J., Burchell, J. Miles, D.W., and Dalziel , (1999) *Bioch. Bioph. Acta.1455,* 301-313.

5. Gendler, S.J., Lancaster, C.A., Taylor-Papadimitriou, J., Duhig, T., Peat, N., Burchell, J., Pemberton, L., Lalani, E.N., and Wilson, D. (1990) *J. Biol. Chem. 265,* 15286-15293.

6. Ligtenberg, M.J.L., Vos, H.L., Gennisen, A.M.C., and Hilkens, J. (1990) *J. Biol. Chem. 265,* 5573-5578.

7. Wreshner, D.H., Haraveuni, M., Tsarfaty, I., Smorodinsky, N., Horev, J., Zaretsky, J., Kotkes, P., Weiss, M., Lathe, R., Dion, A., and Keydar, I. (1990) *Eur. J. Biochem. 189,* 463-473.

8. Haraveuni, M., Tsarfaty, I., Zaretsky, J., Kotkes, P., Horev, J., Zrihan, S., Weiss, M., Green, S., Lathe, R., Keydar, I., and Wreshner, D.H. (1990) *Eur. J. Biochem. 189,* 475-486.

9. Muller, S., Goletz, S., Packer, N., Gooley, A., Lawson, A.M., and Hanisch, F.G. (1997) *J. Biol. Chem.* 272, 24780-24793 (1997).

10. Bennett, E.P., Hassan, H., Mandel, U., Mirgorodskaya, E., Roepstorff, P., Burchell, J., Taylor-Papadimitriou, J., Hollingsworth, M.A., Merkx, G., van Kessel, A.G., Eiberg, H., Steffensen, R., and Clausen, H. (1998) *J. Biol. Chem. 273,* 30472-30481.

11. Wandall, H.H., Hassan, H., Mirgorodskaya, E., Kristensen, A.K., Roepstorff, P., Bennett, E.P., Nielsen, P.A., Hollingsworth, M.A., Burchell, J., Taylor-Papadimitriou, J., and Clausen, H. (1997) *J. Biol. Chem. 272,* 23503-23514.

12. Price, M.R., Rye, P.D., Petrakou, E., Murray, A., Brady, K., Imai, S., Haga, S., Kiyozuka, Y., Schol, D., Meulenbroek, MF, Snijdewint, FG., von Mensdorff-Pouilly, S, Verstraeten, RA, Kenemans, P., Blockzjil, A., Nilsson, K., Nilson, O., Reddish, M., Suresh, MR, Koganty, RR, Fortier, S,

Baronic, L., Berg, A., Longenecker, MB, Hilgers, J, et al. (1998) *Tumour Biol. 19* Suppl 1, 1-20.

13. Rughetti, A., Turchi, V, Ghetti, C.A., Scambia, G., Panici, P.B., Ronucci, G., Mancuso, S., Frati, L., and Nuti, M. (1993) *Cancer Res. 53*, 2457-2459.

14. Kotera, Y., Fontenot, J.D., Pecher, G., Metzgar, R.S., and Finn, O.J. (1994) *Cancer Res. 54*, 2856-2860.

15. Richards, E.R., Devine, P.L., Quin, R.J., Fontenot, J.D., Ward, B.G., and McGuckin, M.A. (1998) *Cancer Immunol. Immunother. 46*, 245-252.

16. Hughes, O.D., Bishop, M.C., Perkins, A.C., Frier, M., Price, M.R., Denton, G., Smith, A., Rutherford, R, and Schibiger, P.A. (1997) *Eur. J. Nucl. Med. 24*, 439-443.

17. Kramer, E.L., Liebes, L., Wasserheit, C., Noz, M.E., Blank, E.W., Zabalegui, A., Melamed, J., Furmanski, P., Peterson, J.A., and Ceriani, R.L.(1998) *Clin. Cancer Res. 4*, 1679-168.

18. Liu. C., Tadayoni, B.M., Bourret, L.A., Mattocks, K.M., Derr, S.M., Widdison, W.C., Kedersha, N.L., Ariniello, P.D., Goldmacher, V.S. Lambert, J.M., Blattler, W.A., and Chari, R.V.(1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 8618-8623.

19. Shinoda, M., Kudo, T., Suzuki, M., Katayose, Y., Sakurai, N, Saeki, H., Kodama, H., Fukuhara, K., Imai, K., Hinoda, Y., and Matsuno, S. (1998) *Cancer Res. 58*, 2838-2843.

20. Pecher, G., & Finn, O.J. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 1699-1704.

21. Graham, R.A., Burchell, J.M., Beverley, P., and Taylor-Papadimitriou, J. (1996) *Int. J. Cancer 65*, 664-670.

22. Apostolopoulos, V, Osinski, C., and McKenzie, I.F. (1998) *Nat. Med. 4*, 315-320.

23. Henderikx, P., Kandilogiannaki, M., Petrarca, C., Von Mensdorff-Pouilly, S., Hilgers, J.H.M., Krambovitis, E., Willen Arends, J., and Hoogenboom, H.R. (1998) *Cancer Res. 58*, 4324-4332.

24. Juweid, M., Neumann, R.D., Palk, C., Perez-Bacete, M.J., Sato, J., van Osdol, W., and Weinstein, J.N. (1992) *Cancer Res. 52*, 5144-5153.

25. Shockley, T.R., Lin, K, Nagy, J.A., Tompkins, R.G., Yarmush, M.L., and Dvorak, H.F. (1992) *Cancer Res. 52*, 367-376 .

26. Saga, T., Neumann, R.D., Heya, T., Sato, J., Kinuya, S., Le, N., Paik, C.H., and Weinstein, J.N. (1995) *Proc. Natl. Acad. Sci. USA 92*, 8999-9003.

27. Fontenot, J.D., Mariappan, S.V.S., Catasti, P., Domenech, N., Finn, O.J., and Gupta, G.(1995) *J. Biomol. Struc. Dynamics 13*, 245-260.

28. Xing, P.-X., Tjandra, J.J., Stacker, S.A., Teh, J.G., McLaughlin, P.J., and McKenzie, I.F.C. (1989) *Immunol. Cell Biol. 67*, 183-195.

29. Xing, P.-X., Reynold, K, Tjandra, J.J., Tang, X.-L., and McKenzie, I.F.C. (1990) *Cancer Res. 50*, 89-96.

30. Geysen, H.M., Meloen, R.H. and Barteling, S.J.(1984) *Proc. Natl. Acad. Sci. U.S.A .81*, 3998-4002.

31. Georges, E., Bradley, G., Gariépy, J. and Ling, V.(1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 152-156.

32. Samelson, L.E. (1991) Chloramine T-catalyzed iodination of soluble proteins. In *Current Protocols in Immunology*. p.8.11.1-8.11.2.

33. Scott, J.K. and Smith, G.P. (1990) *Science 249*, 386-390.

34. Devlin, J.J., Panganiban, L.C. and Devlin, P.E. (1990) *Science 249*, 404-406.

35. Fairbrother, W.J., Christinger, H.W., Cochran, A.G., Fuh, G., Keenan, C.J., Quan, C., Shriver, S.K., Tom, J.Y.K., Wells, J.A., and Cunningham, B.C. (1998) *Biochemistry 37*, 17754-17764.

36. Wiesmann, C., Christinger, H.W., Cochran, A.G., Cunningham, B.C., Fairbrother, W.J., Keenan, C.J., Meng, G., and de Vos, A.M.(1998) *Biochemistry 37*, 17765-17772.

37. Kotera, Y., Fontenot, J.D., Pecher, G., Metzgar, R.S. and Finn, O.J. (1994) *Cancer Res. 54*, 2856-2860.

38. Sudol, M., Chen, H.I., Bougeret, C., Einbond, A, and Bork, P. (1995) *FEBS lett.* 369, 67-71.

39. Aghazadeh, B., and Rosen, M.K. (1999) *Chem Biol. 6*, R241-246.

40. Matsishima, N., Creutz, C.E., and Kretsinger, R.H. (1990) *Proteins Struc. Func.Gen. 7*, 125-155.

41. Bienkiewick, E.A., Moon Woody, A., and Woody, R.W. (2000) *J. Mol. Biol. 297*, 119-133.

42. Litvinov, S.V., and Hilkens, J. (1993) *J. Biol. Chem. 268*, 21364-21371.

43. Dokurno, P., Bates, P.A., Band, H.A., Stewart, L.M.D., Lally, J.M., Burchell, J.M., Taylor-Papadimitriou, J., Snary, D., Sternberg, M.J.E., and Freemont, P.S. (1998) *J. Mol. Biol. 284*, 713-728.

We Claim:

1. An isolated and purified ligand that binds to MUC1 and which comprises the sequence X_1-X_2-X_3 wherein X_1 is Pro, Thr or Ser; X_2 is any amino acid; and X_3 is Pro, Ala, Lys, His, Thr, Asn, Tyr or Phe or an analog or derivative thereof.

2. A ligand according to claim 1 which comprises the sequence X_1-X_2-X_3 wherein X_1 is Pro, Thr or Ser; X_2 is Leu, Ile, Tyr, Ala, Val, Phe, Pro, Met, His, Arg, Asp, Gln or Glu; and X_3 is Pro, Ala, Lys, His, Thr, Asn, Tyr or Phe.

3. A ligand according to claim 2 wherein X_1 is Pro or Ser, X_2 is Arg, Asp, Gln or Phe and X_3 is Pro or Phe.

4. A ligand according to claim 1 wherein X_1 is Pro; X_2 is Arg and X_3 is Phe.

5. A ligand according to claim 4 comprising the sequence FTAPRF (SEQ.ID.NO.:17).

6. A ligand according to claim 4 comprising the sequence PRFPHY (SEQ.ID.NO.:18).

7. A ligand according to claim 4 comprising the sequence TMGFTAPRFPHY (SEQ.ID.NO.:2).

8. A ligand according to claim 1 wherein X_1 is Pro; X_2 is any amino acid and X_3 is Pro.

9. A ligand according to claim 8 comprising the sequence SWWPFP (SEQ.ID.NO.:20).

10. A ligand according to claim 8 comprising the sequence SWWPFPPQPDPA (SEQ.ID.NO.:3).

11. A ligand according to claim 1 wherein X_1 is Ser; X_2 is Asp and X_3 is Pro.

12. A ligand according to claim 11 comprising the sequence SDPLKL (SEQ.ID.NO.:25).

13. A ligand according to claim 11 comprising the sequence YTKSDP (SEQ.ID.NO.:26).

14. A ligand according to claim 11 comprising the sequence YTKSDPLKLLES (SEQ.ID.NO.:4).

15. A ligand according to claim 1 wherein the ligand is selected from the group consisting of TMGFTAPRFPHY (SEQ.ID.NO.:2), SWWPFPPQPDPA (SEQ.ID.NO.:3), YTKSDPLKLLES (SEQ.ID.NO.:4), VVPVHWSRGVVL (SEQ.ID.NO.:5), HIPVAALAPRMT (SEQ.ID.NO.:6), LGLQPPTSALDP (SEQ.ID.NO.:7), TPAFSPLPTDLL (SEQ.ID.NO.:8), ELNTHLATNVFT (SEQ.ID.NO.:9), IDVHSINFLATL (SEQ.ID.NO.:10), THPWSLKSTSFF (SEQ.ID.NO.:11), YITPYAHLAGGN (SEQ.ID.NO.:12), SLPIPSHARLQN (SEQ.ID.NO.:13), YLPYATLSQNSH (SEQ.ID.NO.:14), WHIPPNIGRTFS (SEQ.ID.NO.:15) and TSNPHTRHYYPI (SEQ.ID.NO.:16) or an analog, derivative or fragment of the ligand.

16. An isolated and purified ligand which binds to the sequence PDTRPAPGSTAPPAHGVTSA (SEQ.ID.NO.:1) of MUC1 or a portion thereof.

17. A ligand according to claim 16 which binds to the sequence TRPAPGST (SEQ.ID.NO.:32).

18. A ligand according to claim 16 which binds to the sequence Ala-Pro (SEQ.ID.NO.:28) and/or Pro-Ala (SEQ.ID.NO.:29).

19. A ligand according to claim 16 which binds to the sequence Pro-Ala-X-Gly-X-Thr (SEQ.ID.NO.:30) wherein X is any amino acid.

20. A method of isolating a ligand that binds to MUC1 comprising:

(a) providing a MUC1 target molecule wherein the MUC1 target molecule comprises at least 3 tandem repeats of the sequence PDTRPAPGSTAPPAHGVTSA (SEQ.ID.NO.:1);

(b) contacting the MUC1 target molecule with a peptide or small molecule library, under conditions to allow the MUC1 target to bind to any MUC1 ligands in the library, to prepare a test sample;

(c) panning the test sample under high stringency conditions; and

(d) isolating the MUC1 ligands bound to the MUC1 in the test sample.

21. A method according to claim 20 wherein the library is a peptide phage display library or a synthetic peptide library.

22. A method according to claim 20 wherein the high stringency conditions consist of high salt concentrations.

23. A method according to claim 22 wherein the high salt concentration is about 200-400mM NaCl.

24. A method according to claim 22 wherein the high salt concentration is at least 300 mM NaCl.

25. A method according to claim 20 wherein step (c) comprises of three rounds of panning, the first round at a salt concentration of 75 mM, the second round at a salt concentration of 150 mM and the third round at a salt concentration of at least 300 mM.

26. A method according to claim 21 wherein step (c) comprises four rounds of panning, the first and second rounds at a salt concentration of 75 mM, the third round at a salt concentration of 150 mM and the fourth round at a salt concentration of at least 300 mM.

27. A ligand isolated according to the method as claimed in any one of claims 20 to 26.

28. A use of the ligand according to any one of claims 1 to 19 to detect the presence of MUC1 in a sample.

29. A use of a ligand according to any one of claims 1 to 19 to treat or prevent a cancer associated with MUC1.

30. A use of a ligand according to any one of claims 1 to 19 to prepare a medicament to treat or prevent a cancer associated with MUC1.

31. A use of a ligand according to any one of claims 1 to 19 to prepare a diagnostic agent to detect a cancer associated with MUC1.

32. A method of diagnosing or monitoring a cancer that is associated with MUC1 comprising contacting a sample from a patient with a MUC1 ligand according to any one of claims 1 to 19 and assaying for binding between the MUC1 ligand and MUC1 in the sample, if present, wherein the presence of MUC1 indicates the presence of a cancer that is associated with MUC1.

33. An antibody that binds to a ligand according to any one of claims 1 to 19.

FIGURE 1

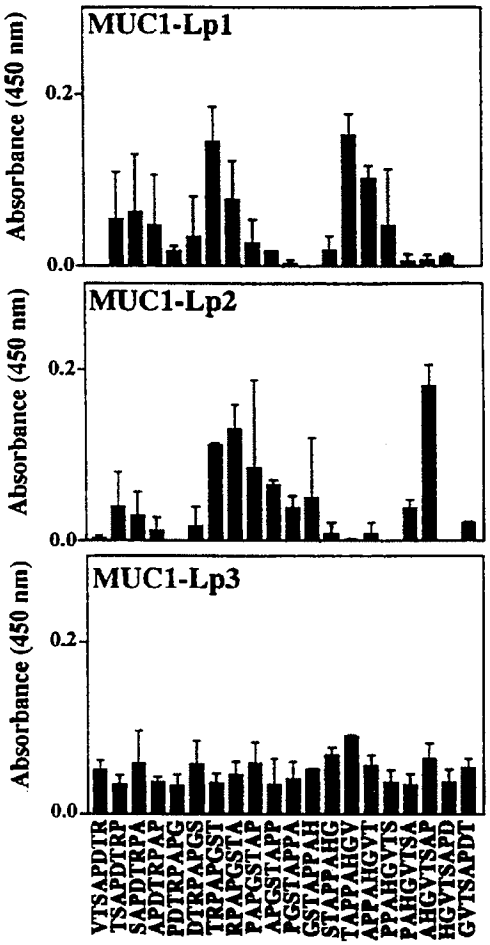

MUC1-Lp1
MUC1-Lp2
MUC1-Lp3
Absorbance (450 nm)
0.2
0.0
VTSAPDTR
TSAPDTRP
SAPDTRPA
APDTRPAP
PDTRPAPG
DTRPAPGS
TRPAPGST
RPAPGSTA
PAPGSTAP
APGSTAPP
PGSTAPPA
GSTAPPAH
STAPPAHG
TAPPAHGV
APPAHGVT
PPAHGVTS
PAHGVTSA
AHGVTSAP
HGVTSAPD
GVTSAPDT
MUC1 octapeptide sequence

FIGURE 2

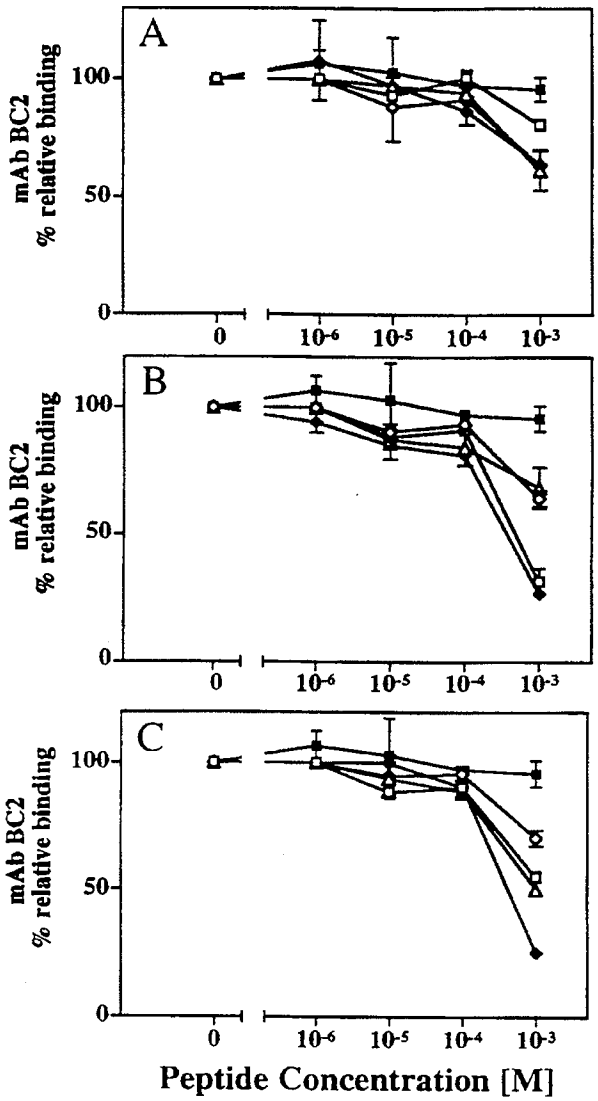

A
B
C
mAb BC2
% relative binding
100
50
0
0
10-6
10-5
10-4
10-3
Peptide Concentration [M]

FIGURE 3

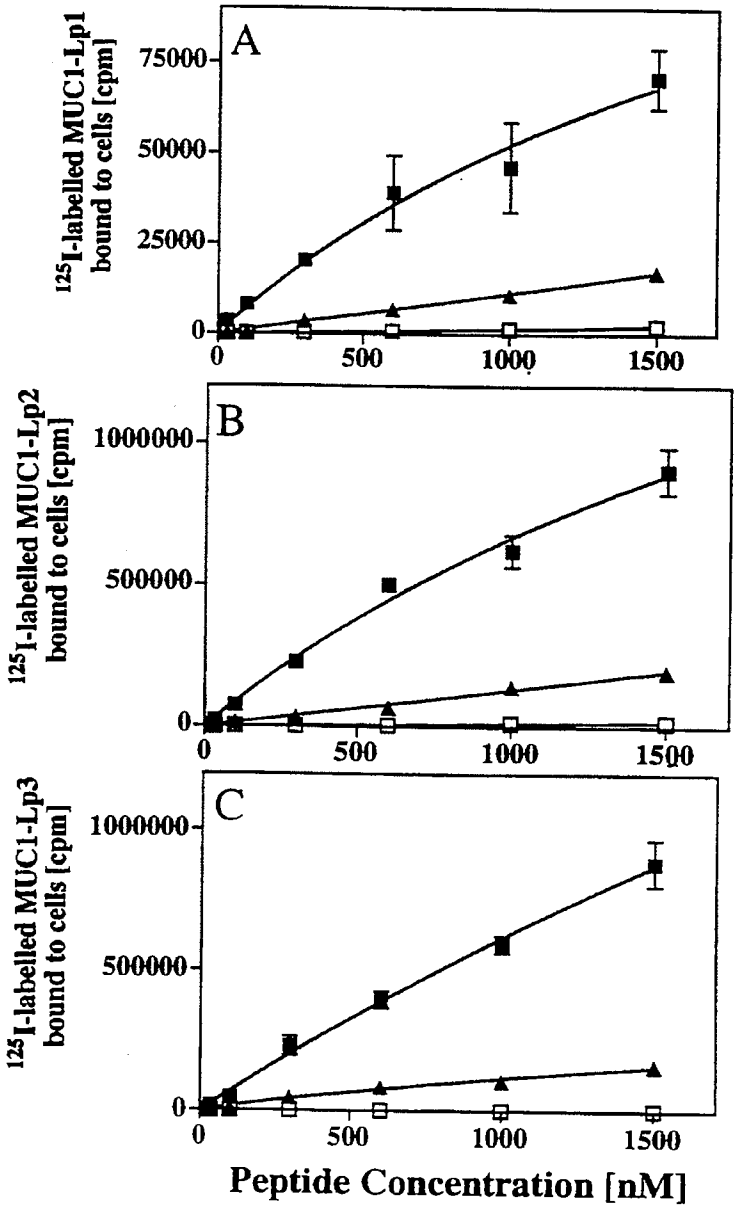
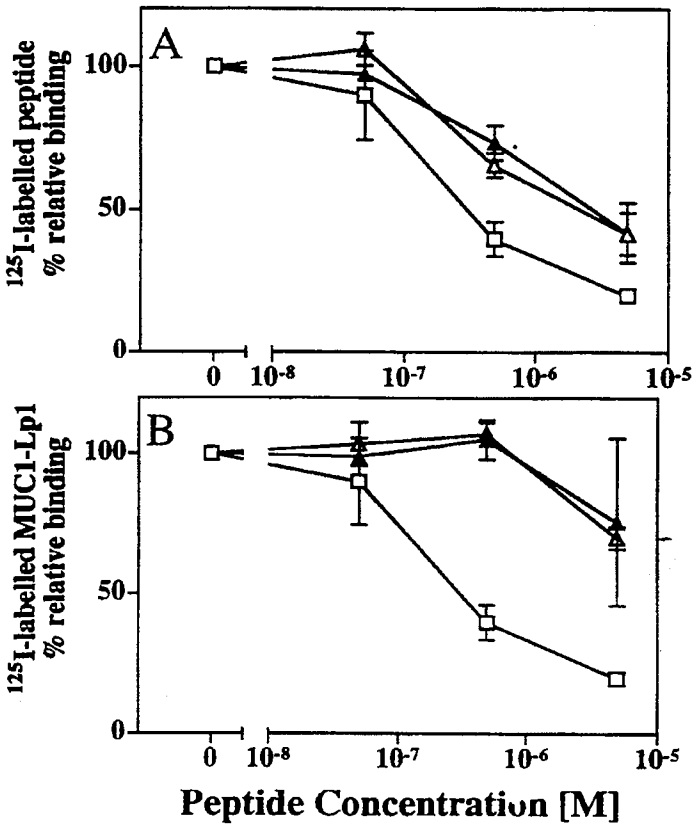

A
B
C
^{125}I-labelled MUC1-Lp1 bound to cells [cpm]
^{125}I-labelled MUC1-Lp2 bound to cells [cpm]
^{125}I-labelled MUC1-Lp3 bound to cells [cpm]
75000
50000
25000
0
1000000
500000
0
500
1000
1500
Peptide Concentration [nM]

FIGURE 4


A
B
^{125}I-labelled peptide % relative binding
^{125}I-labelled MUC1-Lp1 % relative binding
100
50
0
0
10^{-8}
10^{-7}
10^{-6}
10^{-5}
Peptide Concentration [M]

SEQUENCE LISTING

<110> University Health Network
Gariepy, Jean
Yang, Shaoxian

<120> MUC1 Ligand

<130> 10723-13

<140>
<141>

<160> 41

<170> PatentIn Ver. 2.0

<210> 1
<211> 20
<212> PRT
<213> Homo sapiens

<400> 1
Pro Asp Thr Arg Pro Ala Pro Gly Ser Thr Ala Pro Pro Ala His Gly
1 5 10 15

Val Thr Ser Ala
20

<210> 2
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 2
Thr Met Gly Phe Thr Ala Pro Arg Phe Pro His Tyr
1 5 10

<210> 3
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 3
Ser Trp Trp Pro Phe Pro Pro Gln Pro Asp Pro Ala
1 5 10

<210> 4
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 4

Tyr Thr Lys Ser Asp Pro Leu Lys Leu Leu Glu Ser
1 5 10

<210> 5
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 5
Val Val Pro Val His Trp Ser Arg Gly Val Val Leu
1 5 10

<210> 6
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 6
His Ile Pro Val Ala Ala Leu Ala Pro Arg Met Thr
1 5 10

<210> 7
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 7
Leu Gly Leu Gln Pro Pro Thr Ser Ala Leu Asp Pro
1 5 10

<210> 8
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 8
Thr Pro Ala Phe Ser Pro Leu Pro Thr Asp Leu Leu
1 5 10

<210> 9
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 9
Glu Leu Asn Thr His Leu Ala Thr Asn Val Phe Thr

1 5 10

<210> 10
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 10
Ile Asp Val His Ser Ile Asn Phe Leu Ala Thr Leu
1 5 10

<210> 11
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 11
Thr His Pro Trp Ser Leu Lys Ser Thr Ser Phe Phe
1 5 10

<210> 12
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 12
Tyr Ile Thr Pro Tyr Ala His Leu Ala Gly Gly Asn
1 5 10

<210> 13
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 13
Ser Leu Pro Ile Pro Ser His Ala Arg Leu Gln Asn
1 5 10

<210> 14
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 14
Tyr Leu Pro Tyr Ala Thr Leu Ser Gln Asn Ser His
1 5 10

<210> 15
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 15
Trp His Ile Pro Pro Asn Ile Gly Arg Thr Phe Ser
1 5 10

<210> 16
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 16
Thr Ser Asn Pro His Thr Arg His Tyr Tyr Pro Ile
1 5 10

<210> 17
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 17
Phe Thr Ala Pro Arg Phe
1 5

<210> 18
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 18
Pro Arg Phe Pro His Tyr
1 5

<210> 19
<211> 3
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 19
Pro Arg Phe
1

<210> 20
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 20
Ser Trp Trp Pro Phe Pro
1 5

<210> 21
<211> 3
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 21
Pro Xaa Pro
1

<210> 22
<211> 3
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 22
Pro Gln Pro
1

<210> 23
<211> 3
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 23
Pro Asp Pro
1

<210> 24
<211> 3
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 24
Pro Phe Pro
1

<210> 25

<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 25
Ser Asp Pro Leu Lys Leu
1 5

<210> 26
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 26
Tyr Thr Lys Ser Asp Pro
1 5

<210> 27
<211> 3
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 27
Ser Asp Pro
1

<210> 28
<211> 2
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 epitope

<400> 28
Ala Pro
1

<210> 29
<211> 2
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 epitope

<400> 29
Pro Ala
1

<210> 30
<211> 6

<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 epitope

<400> 30
Pro Ala Xaa Gly Xaa Thr
1 5

<210> 31
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 epitope

<400> 31
Ala Pro Asp Thr Arg
1 5

<210> 32
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 epitope

<400> 32
Thr Arg Pro Ala Pro Gly Ser Thr
1 5

<210> 33
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 epitope

<400> 33
Ala Pro Pro Ala His Gly Val Thr
1 5

<210> 34
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 epitope

<400> 34
Arg Pro Ala Pro Gly Ser Thr Ala
1 5

<210> 35
<211> 8
<212> PRT

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 epitope

<400> 35
Pro Ala Pro Gly Ser Thr Ala Pro
1 5

<210> 36
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 epitope

<400> 36
Ala His Gly Val Thr Ser Ala Pro
1 5

<210> 37
<211> 3
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 epitope

<400> 37
Gly Xaa Thr
1

<210> 38
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 epitope

<400> 38
Ala Pro Asp Thr Arg Pro Ala Pro
1 5

<210> 39
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 epitope

<400> 39
Pro Asp Thr Arg
1

<210> 40
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Artificial
Control Peptide

<400> 40
Tyr Thr Val Thr Asn Ser Trp Thr Trp Trp Ser Pro Leu Gln Gln Ala
1 5 10 15

<210> 41
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:MUC1 Ligand

<400> 41
Tyr Ser Trp Trp Pro Phe Pro Pro Gln Pro Asp Pro Ala
1 5 10